

**APPARATUS AND METHOD FOR ATOMIC LAYER  
DEPOSITION ON SUBSTRATES**

Reference to Related Application

[0001] This application claims the priority benefit under 35 U.S.C. §119(e) of provisional application No. 60/431,337, filed December 5, 2002.

Field of the Invention

[0002] The invention relates generally to the field of semiconductor processing and, more particularly, to an apparatus and method for the atomic layer deposition (ALD) of films onto semiconductor substrates.

Background of the Invention

[0003] Various types of reactors are used for atomic layer deposition (ALD) of thin films on semiconductor substrates. During ALD, a semiconductor substrate is exposed to alternating and sequential pulses of at least two mutually reactive reactants. The substrate temperature is chosen to be in a window above the reactants' condensation temperatures and below the reactant's thermal decomposition temperatures, so that during a reactant pulse a monolayer of the reactant adsorbs onto the substrate in a self-saturating manner without condensing or decomposing. After exposing the substrate to a first reactant, the process chamber is purged to remove any excess first reactant from the substrate and to remove any residual first reactant from the gas phase. The substrate is then exposed to a second reactant pulse, the second reactant being reactive with the first reactant.

[0004] Adequate purging and/or evacuating between the alternating reactant pulses is vital for good control of the deposition process. As purging takes time, the ALD process is generally a slow process.

[0005] Because of continual pressures within the semiconductor processing industry for increased throughput, there is a need for methods and apparatuses for ALD that allow efficient purging between subsequent reactant pulses, so that the purge time between

pulses can be reduced and the overall deposition rate of the ALD process can be increased. In addition, because contaminating reactants can remain in a chamber by being deposited onto reactor walls there is also a need for reactors for ALD wherein deposition on the reactor walls is reduced or prevented.

### Summary of the Invention

[0006] In accordance with one preferred embodiment of the invention, a film deposition station is provided for depositing a film onto a substrate. The deposition station comprises a first part and a second part for accommodating a semiconductor substrate between them. The first part and the second part are positioned opposite each other and parallel to a substrate, upon retention of the substrate between the first and second parts. The first part and the second part are also configured to be spaced less than about 2 mm from a main surface of a substrate accommodated between them. At least one of the parts is provided with a heater for heating that part and each part is provided with a set of gas supply channels connected to a source of gas. The source of gas for the first part is configured to supply mutually reactive reactants in a sequence of alternating, separated pulses for atomic layer deposition (ALD).

[0007] In accordance with another preferred embodiment, a reactor is provided for semiconductor processing. The reactor comprises an upper reactor block and a lower reactor block for accommodating a semiconductor substrate therebetween. The upper and the lower reactor blocks are configured to be less than about 2 mm from a major surface of the substrate when the substrate is retained therebetween. Also, the reactor is configured to discharge mutually reactive reactants from at least one of the reactor blocks to the substrate in sequential alternating, separated pulses. The at least one of the reactor blocks comprises a set of gas channels configured to transport and discharge the sequential alternating, separated pulses of reactant to the substrate.

[0008] In accordance with yet another preferred embodiment, a method is provided for depositing a layer on a semiconductor substrate. The method comprises providing an apparatus having a first side section and a second side section located opposite one another. The side sections each have facing planar surfaces and at least one of the side

sections is heated to a temperature higher than about 200°C. The substrate is placed in the apparatus between the first and second side sections and two gas streams, in opposing directions, are applied from the first and second side sections to two opposing planar sides of the semiconductor substrate. The spacing between each of the first and second side sections and the semiconductor substrate is at most about 1 mm and the facing planar surfaces of the side sections extend completely across the opposing planar sides of the semiconductor substrate. At least one of the gas streams provides different reactants in a sequence of alternating, separated pulses for an atomic layer deposition (ALD) process.

[0009] In accordance with another preferred embodiment, a method is provided for semiconductor processing. The method comprises providing a processing apparatus having a first and a second reactor block. A substrate is positioned between the first and the second reactor blocks, wherein the substrate is less than about 2 mm from each of the first and the second reactor blocks after positioning. Mutually reactive reactants are discharged from the first reactor block in alternating, temporally separated pulses onto the substrate. The substrate is heated using the first or second reactor block to one or more desired substrate temperatures, wherein the first reactor block is at a first temperature at which condensation or decomposition of the mutually reactive reactants is substantially prevented.

[0010] In accordance with yet another preferred embodiment, a method is provided for semiconductor processing. The method comprises loading a substrate in a reaction chamber and completely supporting the substrate on a gas cushion in the chamber. While completely supporting the substrate on the gas cushion, mutually reactive reactants are discharged in alternating, temporally separated pulses onto a major surface of the substrate.

[0011] In accordance with another preferred embodiment of the invention, a film deposition station is provided for depositing a film onto a substrate. The deposition station comprises a first part and a second part for accommodating a semiconductor substrate between them. The first part and the second part are positioned opposite each other and parallel to the substrate upon retention of the substrate between the first and second parts. The first part is provided with a first set of gas supply channels connected to a source for a first reactant and a second set of gas supply channels connected to a source for a second reactant. The first and second set of gas supply channels are configured to keep the reactants

separated until discharging the reactants out from the gas supply channels to the substrate, wherein the first and the second reactant are mutually reactive. The deposition station also comprises controls to supply the first and the second reactant from the source for a first reactant and from the source for a second reactant in sequential alternating separated pulses for atomic layer deposition (ALD).

#### Brief Description of the Drawings

[0012] The invention will be better understood from the detailed description of the preferred embodiments and from the appended drawings, which are meant to illustrate and not to limit the invention, and wherein:

[0013] Figure 1 shows a graph of the theoretical reaction regimes of different process recipes for a particular set of ALD reactants in different temperature ranges;

[0014] Figure 2 shows, schematically, a cross-sectional side view of a reactor having a top and a bottom part between which a substrate is accommodated, wherein one reactor part is heated and the other reactor part is cooled, in accordance with certain preferred embodiments of the invention;

[0015] Figure 3 shows, schematically, a cross-sectional side view of the reactor of Figure 2, wherein both reactor parts are heated, in accordance with other preferred embodiments of the invention;

[0016] Figure 4 shows, schematically, a cross-sectional plan view of one of the two reactor parts of Figure 3 as viewed along line A-A and showing gas channels, in accordance with preferred embodiments of the invention;

[0017] Figure 5 shows, schematically, a cross-sectional side view of the reactor part of Figure 4 as viewed along line B-B;

[0018] Figure 6 shows, schematically, a cross-sectional side view of a reactor in which the top reactor part comprises two separated sets of gas supply channels, in accordance with another preferred embodiment of the invention;

[0019] Figure 7 shows, schematically, a cross-sectional side view of the reactor parts of Figure 6, showing a first set of the gas channels in the top reactor part;

[0020] Figure 8 shows, schematically, a cross-sectional side view of the reactor parts of Figure 6, showing a second set of the gas channels in the top reactor part;

[0021] Figure 9 shows, schematically, a cross-sectional side view of a reactor in accordance with yet another preferred embodiment of the invention; and

[0022] Figure 10 shows, schematically, a cross-sectional side view of a more complete reactor module using the reactor of Figure 9, in accordance with another preferred embodiment of the invention.

### Detailed Description of the Invention

[0023] According to preferred embodiments of the invention, a deposition station is provided for depositing films by exposing a substrate to mutually reactive reactants, preferably in alternating pulses, such as for an atomic layer deposition process. The station preferably comprises two parts which delimit a reaction chamber in which the substrate is accommodated. To prevent unintended reaction of the reactants, at least one of the parts is provided with a plurality of gas channels which allow multiple reactants to be flowed through that part without contacting one another until being discharged into the reaction chamber. Preferably, as discussed below, the two parts are positioned close to the substrate during processing to minimize the volume of the reaction chamber and to allow quicker purging and cycling of reactants into the chamber. In addition, as discussed below, at least one of the two parts delimiting the reaction chamber can be heated. Preferably, the temperatures of these two parts are different from the desired process temperature of the substrate and are out of the window of temperatures in which deposition rate is optimized, thus minimizing deposition of reactants on the deposition station surfaces. Advantageously, by minimizing the purge time and the contamination of deposition station surfaces, reactor throughput and the quality of deposited layers are improved.

#### A. Preferred Reactor Module

[0024] While the preferred embodiments can be applied to other reactors known to those of skill in the art, use of a floating substrate reactor is particularly advantageous. In a floating substrate reactor, a substrate can be processed without being mechanically supported;

that is, the substrate can be processed without being directly contacted by a solid support. This enables very uniform and rapid heating of the substrate without the cold spots that can occur in reactors where substrates are mechanically contacted during a semiconductor fabrication process. In addition, parts of the reactor surrounding the substrate are preferably relatively massive such that each has a high heat capacity relative to the substrate, helping to stabilize the temperature of the substrate and minimizing the susceptibility of the reactor to temperature fluctuations upon loading and unloading of the substrate into the reactor. The basic configuration of such a floating substrate reactor is available commercially under the trade name Levitor® from ASM International N.V. of Bilthoven, The Netherlands. In this reactor, a substrate, such as a wafer, is typically accommodated between relatively massive upper and lower blocks. More detailed information about the Levitor® can be found in U.S. Patent No. 6,183,565, the disclosure of which is incorporated herein by reference.

[0025] In addition, a Low Temperature (LT) module of the Levitor®, also available from ASM International N.V. of Bilthoven, The Netherlands, allows for processing at low temperatures and permits the annealing of copper (Cu) films at temperatures in the range of about 200-400°C. A short description of the Levitor® Low Temperature module is provided below. More detailed information about the Levitor® Low Temperature copper anneal module can be found in WO 01/50502 and US Patent Application Publication No. 2002/0092584, the disclosures of which are incorporated herein by reference.

[0026] In some preferred embodiments, both upper and lower reactor blocks of the reactor are provided with heaters, and the temperature of the blocks can be controlled such that the blocks have different temperatures: one block at a higher temperature and the other block at a lower temperature, wherein the desired substrate temperature lies between the high temperature and the low temperature. In other preferred embodiments, both blocks are heated to substantially the same temperature, which is preferably equal to the desired substrate temperature. The substrate temperature can also be regulated by varying the heat transference between the substrate and the reactor blocks. Methods for controlling the substrate temperature by varying heat transference are described in greater detail in co-pending U.S. Applications No. 10/186,269, METHOD AND APPARATUS FOR THE TREATMENT OF SUBSTRATES, filed June 27, 2002 and METHOD FOR THE HEAT

TREATMENT OF SUBSTRATES, filed October 31, 2003, attorney docket No. ASMINT.057AUS, the disclosure of which is incorporated herein by reference.

[0027] Also, in some embodiments, to allow a substrate to be processed at a temperature within a processing window for optimal deposition while the upper and lower blocks are maintained at temperatures outside that window and to be able to quickly cool the substrate, the upper and lower blocks can be set at different temperatures, as noted above. For example, the upper block of the Levitor® Low Temperature module can be kept hot, preferably about 50°C or more above the desired substrate treatment temperature, while the lower block is kept cold, preferably about 50°C or more below the desired substrate treatment temperature or, more preferably, at 40°C or less. The upper block is preferably heated by an array of heating elements, and the lower block is preferably water-cooled. The blocks are provided with a plurality (preferably about 60 to 100) of gas injection channels, spaced from each other and distributed over the surface of the block facing the semiconductor substrate. When the blocks are in a closed position with a substrate loaded between them, the gap between the substrate and each of the blocks is preferably less than about 2.0 mm, more preferably less than about 1.0 mm, and most preferably less than about 0.5 mm. In the illustrated embodiments the gap is about 0.2 mm. Typical powers required to maintain a temperature difference,  $\Delta T$ , of about 400°C between the blocks are in a range of roughly about 5-8 kW for 300 mm semiconductor wafers.

[0028] The temperature of a substrate within the reactor can be varied by switching the type of gas above and below the substrate between a high thermal conductivity gas and a low thermal conductivity gas (*e.g.*, He or N<sub>2</sub>, respectively), and/or by varying the position of the substrate relative to the two blocks (*e.g.*, the relative distance of the substrate from each block). When the blocks are at different temperatures, switching the composition of the gases on one or both sides can change the thermal conductivity of the medium between the substrate and the blocks. Thus, switching the gas compositions on one or both sides of the substrate can be employed to change the substrate temperature within a range between the temperatures of the two blocks. In addition, switching the relative flow rates of gases on either side of the substrate can move the substrate closer to one or the other of the blocks, thereby also changing the substrate temperature.

[0029] When the blocks are in a closed position and arranged close to a substrate, large gas flows (preferably greater than about 5 slm and, in some embodiments, about 10 slm on each side) ensure that the substrate surfaces are effectively protected from the outside ambient. For example, this allows a Cu film to be annealed in the Levitor® Low Temperature module without the need to apply an additional gate valve to prevent oxidation of the Cu film. Because of the large gas flows, the substrate only comes into contact with outside air when the blocks are open and the substrate is relatively non-reactive (*i.e.*, when the substrate is cold in comparison to its temperature during treatment). Preferably, to ensure that the substrate is relatively cold, the substrate temperature is switched to a low value at the end of processing, *e.g.*, by appropriately switching the gases flowing above and below it, before opening the blocks.

[0030] In arrangements where the blocks are maintained at different temperatures, however, the substrate is susceptible to temperature non-uniformities. This is caused by the fact that temperature differences on different sides of a substrate produce a temperature gradient through the thickness of the substrate and thus can cause bowing of the substrate, which can immediately result in a temperature non-uniformity. For example, when the same gas is used on both sides of the substrate, with a  $\Delta T$  between the blocks of about 400°C, and a gap of about 0.2 mm between each block and the substrate, a temperature non-uniformity across the substrate of about 1°C per 1  $\mu\text{m}$  bow can result. Thus, for a typical rest bow of about 10  $\mu\text{m}$ , the difference in temperature of different parts of the substrate can be up to about 10°C. Advantageously, however, this non-uniformity can be reduced when the  $\Delta T$  between the blocks is made smaller or when different gases are used on both sides of the substrate in order to reduce the differences in temperature between the two sides of the substrate.

[0031] Heat-up and cool-down with such an arrangement can advantageously be fast. Using good heat conduction in the thin gas layers results in typical heat-up and cool-down times of preferably less than about 5 seconds and, more preferably, between about 2-3 seconds.

[0032] As noted above, the blocks are preferably relatively massive compared to the substrate. The blocks thus have sufficient heat capacity such that, when heated, heat is



transferred to an unheated substrate loaded between the side sections with negligible temperature loss from the blocks. The blocks preferably are formed of metal and have a preferred vertical thickness of greater than about 10 mm, more preferably greater than about 40 mm.

[0033] It will be appreciated that it is not trivial to deposit films from more than one reactant in reactors such as the Levitor® reactor, as suggested in US patent No. 6,183,565, the disclosure of which is incorporated herein by reference. For example, with conventional CVD processes, it has been found that when deposition gases are pre-mixed, there is a tendency to cause deposition in the gas channels and in the tubing. In addition, there is a tendency toward depletion of reactants in the small volume in between the substrate and the block surface. This depletion can lead to relatively heavy deposition on the substrate immediately below the gas inlet channels, and to less deposition on areas of the substrate in between such inlet points. Consequently, processing using multiple reactants in these reactors can be problematic both in terms of adversely affecting reactor surfaces and in terms of their deposited layers on substrates, which exhibit poor uniformity of both thickness and coverage.

#### B. ALD in the Preferred Reactor Module

[0034] It has been found that the problems that can exist when depositing films using multiple reactants in reactors such as the Levitor® reactor can be avoided when ALD processes are used to deposit the films.

[0035] These problems can advantageously be avoided for a number of reasons. For example, in accordance with some preferred embodiments, when independent gas inlet channels are used for different reactants, there typically will be no deposition in the gas inlet channels.

[0036] Also, when a block facing the substrate device side is kept at a relatively low temperature (but high enough to avoid condensation), there will be substantially reduced ALD growth on the surfaces of that (typically colder) block facing the substrate. Alternatively, when the block facing the substrate device side is kept at a relatively high temperature (but not so high that thermal decomposition of a reactant occurs), such that the

reaction conditions are outside the optimal process window for ALD, substantially reduced ALD deposition will occur on surfaces of this block due to a tendency for reactants to desorb.

[0037] Moreover, even where reactant depletion occurs, self-saturating pulses and true self-limiting reactions do not generally cause non-uniformity in ALD films. This is due to the saturation of absorption sites for reactants, which typically prevents an excess build-up of reactants. Thus, the problems with deposition on reactor surfaces and with poor quality deposited layers can be avoided.

[0038] In addition, various other features of the preferred reactor make it particularly advantageous for ALD processes. For example, the reactor volume is particularly small, due to the small spacing of the substrate from each of the upper and lower reactor blocks of about 2 mm or less, preferably about 1 mm or less and, more preferably, of about 0.2 mm or less. The small reactor volume allows a rapid switching between the alternating reactant pulses in ALD; any traces of a reactant pulse still residing in the reactor after termination of the pulse can be purged from the reactor in a very short time.

[0039] Not only is the reaction space volume small, but the gas distribution system used with the preferred reactor, described in further detail below, also has a small volume. Unlike conventional showerheads, the gas distribution system does not comprise a relatively voluminous gas distribution section to distribute gas laterally and uniformly over the substrate. Instead, the gas distribution system preferably comprises a set of horizontal channels to distribute the gas in a lateral direction over the substrate, which is typically placed in the reactor in a horizontal position. Preferably, a set of vertical channels, each vertical channel preferably in communication at one end with one of the horizontal channels and at the other end with the reaction space, discharges the gas into the reaction space. The channels are preferably drilled in the reactor blocks, which are preferably massive plates. The horizontal channels are relatively small in diameter, preferably about 3-5 mm and more preferably about 4 mm in diameter. The vertical channels are preferably about 1-3 mm in diameter and more preferably about 2 mm in diameter. Consequently, the total volume of the set of horizontal and vertical channels is small and easily purged in comparison to a conventional showerhead that comprises a relatively open gas distribution volume upstream

of a perforated faceplate. In addition, the small gas volumes at both sides of the substrate further ensure that the total reactor volume is purged particularly efficiently.

[0040] Another advantage of the preferred reactor for ALD is the ability to control the temperature of the substrate at a value different from the temperature of the reactor blocks, thereby helping to minimize deposition on the reactor blocks. It will be appreciated that true self-limited atomic layer deposition will only occur in a certain temperature window. Preferably, the reactor block from which the reactants are dispersed is maintained at a temperature outside the window for optimized ALD growth, and preferably at a temperature where little or no growth occurs, so that little or no deposition occurs on this block, while the substrate is maintained at a temperature within the window for ALD growth.

[0041] An additional advantage of the preferred reactor, *e.g.*, the Levitor® LT module is that, during processing, the temperature of the substrate can rapidly be switched. This can beneficially be exploited in a manner as disclosed by Kurabayashi *et. al.* in a paper entitled "Temperature synchronized molecular layer epitaxy," *Applied Surface Science* 82/83 (1994) 97-102, the disclosure of which is incorporated herein by reference. Kurabayashi describes, among other examples, a process wherein a GaAs film is deposited on a substrate by exposing the substrate sequentially and repeatedly to pulses of AsH<sub>3</sub> and TMG (Tri Methyl Gallium), and wherein the substrate temperature is 500°C during the AsH<sub>3</sub> pulse and 450°C during the TMG pulse. The higher temperature during the AsH<sub>3</sub> pulse is preferred for achieving a sufficient reaction rate, whereas the lower temperature during the TMG pulse is preferred for avoiding premature thermal decomposition of the TMG. In general, this technique allows the use of temperature sensitive precursors during the low temperature part of the cycle whereas the high temperature part of the cycle ensures that a sufficient reaction rate and/or film quality is achieved. This is beneficial in the deposition of metals, metal oxides, and metal nitrides.

[0042] In some situations, cycling the temperature is a good alternative to the use of a plasma during the deposition process. Further, when depositing multicomponent materials, temperature cycling allows the use of combinations of reactants, in particular metal compounds, whose respective process windows do not match.

[0043] Advantageously, the preferred reactor is well suited for rapid switching of substrate temperature without requiring a large peak power and large variations in power for the heating system. It will be recognized that the temperatures between which processing is switched can be tailored depending on the reactants used. Also, the timing of the temperature switching can be tailored with respect to the timing of the gas pulses, to achieve optimum results. After supplying a first reactant at a first substrate temperature, the temperature of the substrate can be switched to the second temperature at the beginning, during or after the supply of the second reactant pulse. Although the moment of temperature switching can be shifted in time with respect to the moment of reactant switching, it will be appreciated that the period in which the reactor is at the first temperature should overlap with presence of the first reactant and the period in which the reactor is at the second temperature should coincide with the presence of the second reactant. Preferably, the duration of the period between which the temperature and reactants are switched is the same and switching of the temperature and reactants occurs at substantially the same frequency.

C. Detailed Description of the Figures.

[0044] Reference will now be made to the Figures, wherein like numerals refer to like parts throughout.

[0045] With reference to Figure 1, the temperature window for optimal ALD is explained in further detail. In this figure, the growth of the film per cycle is given as a function of temperature. The horizontal axis indicates the deposition temperature and the vertical axis indicates the amount of film growth per deposition cycle. W denotes the temperature window in which optimized atomic layer deposition occurs. In an ideal case for the temperature window W, represented in the drawings as a horizontal line across the vertical growth axis, one monolayer of film is deposited per full cycle. It will be appreciated, however, that due to surface reconstruction or steric hindrance by large surface ligands, less than a monolayer per cycle is typically deposited in an actual ALD process performed within this window, but optimal self-limited growth still obtains.

[0046] Outside the window, various factors can cause greater than or less than optimized growth per cycle. For example, at the low temperatures at the low end of the

temperature window (W), depending on whether a reactant condenses at those temperatures, more or less than a monolayer can be deposited.  $L_1$  denotes a region of increased growth per cycle, exhibited by reactants that condense at those temperatures.  $L_2$  denotes a region of decreased growth per cycle, for other reactant combinations. The temperature dependence described by  $L_2$  is indicative of a process with combinations of reactants for which the reaction is activation energy limited, where the reactivity of the reactants becomes too low. It should be noted, however, that even for chemistries exhibiting a low-temperature curve like  $L_2$ , increased deposition can eventually occur at lower temperatures than those shown, since condensation usually will take place at a low enough temperature.

[0047] On the other hand, at temperatures at the high end of the temperature window, other factors can lead to deposition of greater than or less than one monolayer per cycle.  $H_1$  indicates a situation where the growth per cycle is above the optimized, self-limited growth. This can occur for some process recipes when the temperature is so high that thermal decomposition of one of the reactants occurs or where non-volatile reaction by-products are formed.  $H_2$  represents a situation where the growth per cycle is less than one monolayer. This may be the result of desorption or dissociation of a surface ligand that is needed to activate the surface for the next reactant. It should be noted, however, that even for chemistries exhibiting a high-temperature curve like  $H_2$ , increased deposition can eventually occur at even higher temperatures than those shown, when thermal decomposition causes deposition of more than one monolayer per cycle.

[0048] As shown in Figure 1, maintaining the substrate surface temperature within a certain window is important for achieving the desired amount of growth of a thin film on a substrate by ALD. Preferably, information about the growth curve for a particular reaction is used to select an appropriate substrate temperature. Thus, a substrate temperature is preferably chosen that falls within the window in which ALD occurs. Advantageously, the preferred embodiments also take advantage of the other information available from the growth curves to minimize deposition on the reactor walls and to provide an improved ALD method and apparatus.

[0049] Exemplary embodiments will now be described with reference to Figures 2-9. Figure 2 shows schematically a reactor 100 having the general features of a Levitor®

Low Temperature module. The reactor 100 comprises an upper reactor block 120 and a lower reactor block 160, between which a substrate or wafer 110 is accommodated. It will be appreciated that the surfaces of the blocks 120 and 160 facing each other delimit a reaction chamber for processing the wafer 110. The upper reactor block 120 is preferably heated through resistance heating, using an array of heating elements 122, each heating element connected with wires 123 to a source of electrical energy. Gas flows through the upper block 120 via a gas inlet opening 124, and gas inlet channels 125 and 126. Then the gas is dispersed over an entire upper or first major surface 110a of the wafer 110 via a plurality of gas dispersion channels 128 in communication with the gas inlet channel 126. Finally, the gas is discharged from the upper block 120 towards a first surface 110a through gas injection channels 129.

[0050] Shown on the other side of the wafer 110, the lower reactor block 160 is provided with cooling channels 162 for the passage of a cooling fluid. Gas flows through the lower block 160 via a gas inlet opening 164 and gas inlet channels 165 and 166. Then the gas is preferably dispersed over an entire wafer lower or second surface 110b via a plurality of gas dispersion channels 168 in communication with the gas inlet channel 166. Finally, the gas is discharged from the lower block 160 towards the second major surface 110b of wafer 110, opposite the first surface 110a, through gas injection channels 169. In the reaction space between the blocks 120 and 160, the gases are preferably flowed in a radial direction and are exhausted from the reaction space via a peripheral gap 130 between the blocks 120 and 160.

[0051] It will be appreciated that the gas dispersion channels 128, 168 can also be provided in individual sections which do not extend over the whole of the surfaces 110a and 110b, respectively. For example, while one horizontal gas inlet channel 125, 165 and one vertical gas inlet channel 126, 166 are illustrated distributing gas over the entire surface 110a or 110b, respectively, more than one gas inlet channel 125, 165 and 126, 166 can be provided, each providing gas to a gas dispersion channel that only extends over a limited section of the surface 110a or 110b.

[0052] It will also be appreciated that a lower block need not be actively cooled. For example, in another embodiment, as shown in Figure 3, the lower block 150 is provided with the heating elements 152, instead of cooling fluid channels, to maintain the lower block

150 at a raised temperature. The temperature of the lower block 150 can be the same as the temperature of the upper block 120 or it can be a different temperature. Preferably, the heating elements 152 are resistive heaters and are each connected via wires 153 to a source of electrical energy. Gas is flowed through the lower block 150 via a gas inlet opening 154 and gas inlet channels 155 and 156. Then the gas is dispersed over the entire wafer surface 110b via a plurality of gas dispersion channels 158 in communication with the vertical gas inlet channel 156. Finally, the gas is discharged from the lower block 150 towards the second surface 110b of the wafer 110, opposite the first surface 110a, through gas injection channels 159.

[0053] The configuration of the gas channels in the blocks is shown in greater detail in Figures 4 and 5. Figure 4 shows a horizontal cross-section of a block 420 similar to the block 120 of Figure 3, taken across plane A-A of Figure 3. Like elements are referenced by reference numerals having like last two digits relative to the elements of Figure 3. The block 420 is provided with the gas dispersion channels 428 in a radial configuration. Each gas dispersion channel 428 is provided with a plurality of gas injection channels 429.

[0054] Figure 5 shows a vertical cross-section taken along plane B-B in Figure 4. Gas enters the block 420 through the inlet 424, which is in communication with the gas inlet channel 425, through which gas is discharged into the annular vertical inlet channel 426. The gas dispersion channels 428 are in communication with the vertical channel 426. The annular channel is formed by inserting a plug 434 into a cylindrical hole 433 in the block 420. As the plug 434 has, over a part of its length, a smaller diameter than the cylindrical hole 433, the vertical inlet channel 426 is left open. The gas dispersion channels 428 are preferably drilled in the block 420 and their radially outward ends are preferably closed by inserting plugs 432.

[0055] Preferably, the cross-sectional area of each of the gas dispersion channels 428, as represented by, *e.g.*, their diameter(s), is larger than the cross-sectional area of each of the gas injection channels 429 (Figure 4), so that a roughly homogeneous distribution of gas through the gas dispersion channels 428 and vertical gas injection channels 429 is achieved. In a preferred embodiment, the diameter of the gas dispersion channels 428 is about 3-5 mm

and more preferably about 4 mm and the diameter of the gas injection channels 429 are about 1-3 mm and more preferably about 2 mm.

[0056] In addition, the gas injection channels 429 themselves preferably comprise a first section with a relatively small first diameter, directly connected to the gas dispersion channels 428, and a second section with a wider second diameter, which opens to the reaction space occupied by the substrate 110 (Figures 2 and 3). Preferably, the first diameter is about 1 mm or less and the second diameter is about 1 mm or more. In one preferred embodiment, the first diameter is about 0.25 mm and the second diameter is about 2 mm, the length of the first section is about 1 mm and the first section is at an upstream end intersecting with a gas dispersion channel 428 whereas the length of the second section is about 3 mm. The downstream end of the first section is in communication with the upstream end of the second section. The downstream end of the second section opens into the reaction space. In this way, a narrow restriction in gas injection channel 429 is provided, whereas the second wider section allows expansion of the gas jet and reduction of the gas velocity before discharging the gas into the reaction space, so that the occurrence of gas jet marks on the substrate is advantageously avoided.

[0057] It will be appreciated that various other arrangements are possible for the gas channels, so long as they allow for gas to be distributed over a substrate surface. For example, the gas dispersion channels 428 need not be in a radial configuration but, *e.g.*, can be arranged according to a pattern of parallel, equidistant lines, connected at one end with a gas inlet channel oriented perpendicular to the gas dispersion channels, forming a rake type configuration. Further, the two sets of separated gas dispersion channels need not necessarily be vertically displaced but they can be, *e.g.*, two oppositely oriented, interlaced rakes, disposed at about the same vertical level. Similarly, it will also be appreciated that various other arrangements are possible for the gas inlet 424 and the gas inlet channels 425 and 426, so long as they allow gas to be provided to the gas dispersion channels 428.

[0058] With reference to Figure 6, a particularly advantageous embodiment is shown, wherein the upper block 120 is provided with two different, separated sets of gas distribution channels, one set of channels for each one of the two mutually reactive reactants typically used during the deposition process. The first set comprises a gas inlet opening 124,



in communication with one end of a gas inlet channel 125, the other end of the gas inlet channel 125 in communication with a vertical inlet channel 126, discharging into an annular gas distribution channel 127. A plurality of gas dispersion channels 128 is in communication with the annular gas distribution channel 127. Finally, gas is discharged from the gas dispersing channels 128 through gas injection channels 129 towards the first surface 110a of the wafer 110.

[0059] A second set of gas distribution channels, separated from the first set, is configured similarly to the first set. The second set comprises a second gas inlet opening 144, in communication with one end of a second gas inlet channel 145, the other end of the second gas inlet channel 145 being in communication with a second vertical inlet channel 146, discharging into a second annular gas distribution channel 147. A plurality of second gas dispersion channels 148 are in communication with the annular gas distribution channel 147. Finally, gas is discharged from the gas dispersing channels 148 through the second gas injection channels 149 towards the first surface 110a of wafer 110. Preferably, other features in Figure 6 are similar as those in Figure 3.

[0060] It will be appreciated that Figure 6 is a schematic, cross-sectional side view, showing both sets of channels in the upper block 120 simultaneously. Preferably, the two sets of horizontally extending gas dispersion channels 128 and 148, shown with the first set at a first vertical level and the second set at a second vertical level, are not just vertically displaced with respect to each other, but are also horizontally displaced. In this way the vertical gas injection channels 149 communicating with the upper set of gas dispersion channels 148 do not intersect the lower set of gas dispersion channels 128.

[0061] An exemplary such arrangement is illustrated in greater detail in Figures 7 and 8. In Figure 7, which is a cross-sectional view taken through a first plane, the first set of gas channels 124 – 129 is shown. Also shown is the annular gas distribution channel 147, which is an annular ring and visible at any cross-section taken through the center axis of the reactor blocks 120 and 150. In Figure 8, which is a cross-sectional view taken through a second plane, slightly rotated with respect to the first plane, the second set of gas channels 144 – 149 is shown. Also shown is the annular gas distribution channel 127. It will be

appreciated that while no channels are illustrated in the lower block 150 for clarity of illustration, the lower block 150 can be provided with channels, as discussed herein.

[0062] It will be appreciated that the number of sets of channels such as the gas channels 124 – 129 and 144 – 149 is not limited to two; rather, the number of separated sets of channels is only limited by the number of set of channels that can be accommodated in the block 120. For example, in other arrangements, three or more separated sets of channels can be accommodated in the block 120. Also, the lower block 150 can be provided with separated sets of channels, which can be arranged differently or can mirror those of the upper block 120.

[0063] A further embodiment is shown in Figure 9. The gas inlets 124 and 144 of the upper block 120 are now provided at the top and through the center of the upper block 120 and the gas inlet 154 of the lower block 150 is provided at the bottom and through the center of the lower block 150. It will be appreciated that the gas inlets 124 and 144 can be connected to gas lines 224 and 244, which are in turn connected to one or more controllers 250, which preferably comprise separate valves commonly controlled by a computer programmed to provide gases to the gas inlets 124 and 144 in sequential alternating, separated pulses. The controller 250 is preferably also connected to reactant sources 324 and 344. In addition, while illustrated only in Figure 9 for ease of description, it will be appreciated that each of the gas inlets described above with reference to the other Figures can also be connected to gas sources and controllers for regulating the flow of gas from those gas sources. Also, the controller can be configured for inert gas valve control as described below.

[0064] Figure 10 is a schematic cross section of an ALD module based on Levitor® principles in accordance with another preferred embodiment. The configuration of the blocks 120 and 150 is similar to that shown in Figure 9. A chamber around the blocks is delimited by an upper wall 190 and a lower wall 192. The upper block 120 is stationary and in fixed connection with the upper wall 190. The lower block 150 is vertically movable to allow for loading and unloading of the wafer 110 in between and out from between the blocks 120 and 150 through an opening 195. The mechanism for vertically moving the lower block 150 is schematically indicated by 194. A removable shield 193 is preferably connected

to the lower block 150. Reaction gases are exhausted from the reaction space between the blocks via the annular channel 196.

[0065] In a preferred embodiment, the upper block 120 is maintained at a relatively low temperature, below the desired substrate temperature, and the lower block 150 is kept at a relatively high temperature, above the desired substrate temperature. Preferably, the removable heat shield 193 is kept at the same (relatively high) temperature as the lower block 150. As the shield 193 extends inwards such that its inside edge falls underneath the wafer 110 and reactant not adsorbed or reacted on the wafer 110 flows off the top surface 110a of the wafer and down to the shield 193, atomic layer deposition on the process system surfaces is concentrated on this shield 193, and not on other parts of the system. Advantageously, by removing and replacing the shield, *e.g.*, once every 10-20  $\mu\text{m}$  of film deposition, a low particle generation level can be obtained. In addition, by making the shield 193 longer, so that it extends further downwardly into the annular channel 196, it can also fulfill the role of largely depleting the outgoing gas stream of reactants, in that way cleaning up the gas that is exhausted.

[0066] It will be appreciated that the system for supplying gas to the reactor 100 can be configured in various ways known to those of skill in the art. Typically, the gas supply system will comprise one or more controllers which are programmed to provide separate, alternating gas pulses to the reactor 100. In some arrangements, each of the blocks of a Levitor® reactor can have separate gas inlets which can be connected to separate gas feeds, as described in U.S. Patent No. 6,183,565, the disclosure of which is incorporated herein by reference, or as shown in the Figures 6 and 9 discussed herein. In such a case, it will be appreciated that a gas switching mechanism is preferably provided upstream to allow for an ALD sequence of alternating and repeating reactant pulses, separated by reactant removal (*e.g.*, purging) steps. An exemplary switching arrangement is sometimes referred to as “inert gas valving,” and is described in U.S. Patent Application Publication US 2001/0054377, the disclosure of which is incorporated herein by reference. In such an arrangement, pulses of reactant vapor are directed from the reactant source container towards the reaction chamber and, through switching of an inert gas flow, the reactant vapor flow is alternately: (i) directed to the reaction chamber by an inert gas flow from the source

container towards the reaction chamber and then (ii) prevented from flowing from the source container to the reaction chamber by an inert gas flow in a reverse direction in a part of the conduit connecting the source container and the reaction chamber. The inert gas valving system is most particularly useful with liquid reactants that are vaporized prior to being employed in deposition. In other embodiments, the switching system can comprise conventional valves and valve control systems, particularly when more volatile gas sources are employed.

[0067] In the arrangements described above having separate gas inlets and separated sets of gas distribution channels in a block, each inlet can be connected to a different gas feed. For example, the separate channels can comprise a “spoke” arrangement, in which different spokes are connected to different reactant gas feeds.

[0068] In the preferred embodiments, deposition of a film can be performed by maintaining a  $\Delta T$  of, *e.g.*, about 100°C or more and, more preferably, about 200°C or more between the blocks. In the reactor 100 of Figure 2, the lower block 160 is cooled and can be at a temperature below about 40°C. The upper block 120 can be heated to a temperature of about 350°C. By the appropriate choice of the types of gases and gas flow ranges below and above the wafer 110, a wafer temperature between that of the upper and lower blocks 120, 160, *e.g.*, of about 300°C, can be achieved. This way of operation is beneficial for processes for which the increased temperature of about 350°C results in desorption of the reactants from the high temperature surface but still does not result in substantial thermal decomposition of the reactants. In other words, with the block at about 350°C, the process should be in the H<sub>2</sub> regime discussed with reference to Figure 1. In such an arrangement, it will be appreciated that the temperatures of the reactor blocks are preferably chosen based on the physical properties of the reactants, *e.g.*, the decomposition and/or condensation points of the reactants.

[0069] In other embodiments, the lower upper block 150 can be heated, rather than cooled, but still maintained at a temperature that is lower than the desired wafer temperature, while the upper block 120 (Figures 3 and 6-10) is maintained at a temperature that is higher than the desired wafer temperature.

[0070] In yet another embodiment, the upper block 120 can be maintained at a lower temperature than the desired wafer temperature. When the desired wafer temperature is, *e.g.*, about 300°C, one can keep the upper block 120 at about 230°C and the lower block 150 at about 330°C. As noted above, this can be realized by the appropriate choice of the types of gases and flows below and above the wafer 110. For example, because the temperature of the lower block 150 is relatively close to the desired wafer temperature, one possibility is to physically keep the wafer 110 close to the lower block 150 to heat the wafer 110 to the desired wafer temperature, which is closer to the lower block temperature than it is to the upper block temperature. The temperature of the upper block 120 is preferably not be so low that condensation of the reactant occurs on it. In other words, for the block at about 230°C, the process is not in the  $L_1$  regime, as discussed above regarding Figure 1. Preferably, for the upper block 120 at about 230°C, the process is in the  $L_2$  regime.

[0071] Processing according to the embodiments of the invention offers numerous advantages. For example, for embodiments employing separated channels for each reactant, by eliminating pre-mixing of the gases in the gas introduction manifold up to the point where the gases enter the deposition volume, deposition in the gas channels leading to the reaction chamber is substantially prevented. In addition, a small nitrogen purge is preferably constantly flowed through the gas channels to ensure that the gas inlet system remains clean up to the final injection point facing the wafer.

[0072] By preferably keeping the temperature of the upper block outside the window for optimal ALD growth, but outside the range in which thermal decomposition can occur and outside the range in which condensation can occur, deposition on the upper block is also substantially prevented. To prevent thermal decomposition or condensation, the temperature of the upper block can be kept sufficiently below (*e.g.*, about 70°C) or above, respectively, that of the wafer. As noted above, in addition to the wafer process temperature, the selection of temperatures also depends on the reactants used. Similarly, the temperature of the lower block is preferably also outside of the window for optimal ALD growth. Note that all the embodiments described herein can simply be inverted so that the reactants flow through the lower block, or reactants can flow through both blocks for double-sided deposition.

[0073] The inert gas that preferably flows through the gap in between the wafer and the lower block substantially prevents deposition on the backside of the wafer, or on the lower block surface facing the wafer. Calculations indicate that the reactants present in the flow above the wafer typically will not penetrate the narrow space between the wafer and the lower block. This is because the flow rate below the wafer is relatively large (greater than about 5 slm) and the gap narrow (about 0.2 mm). As the Reynolds number is very small, this flow is expected to be substantially laminar.

[0074] As none of the reactants typically enter the volume between the wafer and the lower block, back feeding of reactants from this volume to the volume between the wafer and the upper block, into which reactants are introduced, can be substantially eliminated. Consequently, by preventing the different reactant gases from coming into contact with one another, CVD growth can be prevented.

[0075] The preferred embodiments advantageously allow reactants to be switched quickly. In principle, gas switching is much faster at low pressures than under atmospheric conditions. This has been a primary reason why conventional ALD reactors operate at reduced pressures. However, the extremely small reaction volume in a reactor according to the preferred embodiments, such as a Levitor® reactor, in combination with relatively large gas flows, ensures a high gas velocity, even at atmospheric pressure. For example, when a flow of, *e.g.*, about 10 slm is distributed in the volume above the wafer, and the gap is about 0.2 mm, the gas velocity at the edge of a 300 mm wafer is about 0.9 m/s at room temperature. More generally, the gas velocity is preferably arranged to be about 0.1 m/s to 10 m/s. Under the preferred conditions, a typical distance of the wafer radius, *i.e.*, about 0.15 m, is traversed in about 0.02 seconds to about 2 seconds, or about 0.17 seconds in this particular example. At temperatures above room temperature this time will even be shorter due to thermal expansion of the gases. Thus, the reactors can be quickly cleared of reactants for the next reactant pulse, regardless of whether separated gas channels are provided for the reactants.

[0076] In some embodiments, rotation (preferably about 20 rpm) of the wafer can be introduced by flowing gas through a number of additional gas channels. Arrangements for such rotation are known to those of skill in the art and was accomplished, for example, in the first version of the Levitor® 4000 reactor, produced by ASM International N.V. of

Bilthoven, The Netherlands. Such rotation can improve uniformity by more evenly exposing the wafer to reactants. Also, the uniformity can be further improved by the fact that the wafer tends to be self-centering in such arrangements, positioning itself at the center of the reactor.

[0077] In addition, in reactors according to the preferred embodiments, the heat-up rate is fast, preferably on the order of a few seconds. In contrast, in conventional systems this can take several 10's of seconds. Thus, the fast heat-up rate minimizes the time needed for the wafer to reach the equilibrium temperature.

[0078] Also, particle generation is typically low. This is because the preferred large flow above the wafer prevents foreign material from penetrating the space above the wafer. Moreover, by preventing deposition on the blocks facing a wafer, build up of films on that surface is prevented, thus preventing formation of particles.

[0079] The cost of processing according to the preferred embodiments is advantageously low. The main reasons for this are that no vacuum is required, and, given the relatively low temperatures involved, easy to machine materials (*e.g.*, aluminum) can be used for the reactor.

[0080] The footprint of the preferred reactors is also advantageously small. For example, the current Levitor® LT module has a width of about 45 cm and a depth of about 75 cm. In addition, two LT modules can be placed above one another, further minimizing the area occupied by the reactors.

[0081] In an advantageously simple embodiment of the invention, the apparatus described above is operated at atmospheric pressure, so that the use of vacuum pumps can be omitted. It is not necessary, however, to operate the apparatus at atmospheric pressure. A pressure of about 10 mbar to about 100 mbar would still be high enough to have a wafer supported by a gas cushion. In addition, an advantage of reducing the pressure is that it results in an expansion of the gas flows and a more efficient purging of the reaction space.

[0082] It will be appreciated that while discussed with reference to a Levitor® reactor for ease of description, the preferred embodiments are not limited to such a reactor. For example, the gas distribution system and the heating arrangement described above can readily be applied to other reactors to similar advantageous effect.

[0083] It will also be appreciated that although the invention has been described in terms of certain preferred embodiments, various other combinations, omissions, substitutions and modifications can be made to the embodiments described above without departing from the scope of the invention. All such modifications and changes are intended to fall within the scope of the invention. Accordingly, the present invention is not intended to be limited by the descriptions of the preferred embodiments, but instead is to be defined by referenced to the appended claims.